

PAD 6250 Data Paper Assignment

Most academic and policy research focuses on explaining variation in a dependent variable, asking what causes some people to act or believe one way and others to act or believe differently. For example, what demographic characteristics and attitudes of respondents influence their salaries, political tolerance, political ideology, or opinions on government spending? In this paper, you will choose a research question that interests you, develop several hypotheses, use an appropriate dataset to test those hypotheses, present your findings, and conclude whether your hypotheses are correct or not.

1. The first step is to choose an issue/topic that interests you. This assignment is easier to complete if you pick something interesting. Then develop that into a research question (that is, what you're trying to figure out with your research project). If you're interested in public opinions on a topic, for example, you might be researching what individual characteristics would influence opinions on that issue. Although we often use public opinion examples in classes, there are many other kinds of topics that are frequently more interesting for public and nonprofit administrators. Again, try to follow your interests.
2. Go to the Roper Center or the ICPSR or some other data source and find a data set that includes the necessary variables/questions. (Look at the Roper Center guide on D2L for instructions on how to use that site.) You might have access to data through your job or internship, which is probably fine to use (but please discuss with me). Alternatively, you may use either the "WashingtonPostKaiser-2015May(birth control)" dataset or the "Pew-2015Jul (same-sex marriage)" dataset on the D2L site. *However, the best papers (that is, those that receive A grades) will investigate an original question with a new dataset, instead of using one of these two surveys.*
3. Be sure you download a codebook if you are working with someone else's data so that you can double-check that you understand how the variables are coded. If you use the birth control dataset, go to the Datasets folder in D2L and open the "WashingtonPostKaiser-2015May(birth control)" dataset and the "WashingtonPostKaiser-2015May(birth control)" codebook. Use "Utilities → Variables" or "Analyze → Descriptive Statistics → Frequencies" in SPSS (or read the codebook) to explore the variables and make sure you understand how the variables that interest you are coded. You should also check the exact question wording if this is a survey.
4. Choose an interesting dependent variable. The most obvious one in the birth control dataset is whether the respondent thinks insurance plans should be required to cover the cost of birth control, but there are other possibilities.
5. Choose three (or more) independent variables that you expect to have a causal impact on your dependent variable. For some topics, it might be appropriate to use a single independent variable and multiple dependent variables; please talk to me if you think this might be appropriate for you.
6. You may need to edit some of the existing variables or recode them into new variables. (See past homework or the guide on D2L on how to recode variables, change missing values, and insert variable and value labels.) Remember that we can only make crosstabs with nominal, ordinal, and dummy variables.

- a. If one of your ordinal or nominal variables has several values, you might need to recode it to combine some values.
 - b. If you want to use an interval-level variable, you'll need to recode it into an ordinal or dummy variable.
 - c. If you want to exclude some values of a variable, you'll need to set those values as missing in SPSS.
7. Use crosstabs to test your hypotheses. Have SPSS calculate appropriate percentages, gammas, and chi-square statistics. Write a short report using this format:
 - Title
 - Research Question:
 - Write 2-4 sentences making clear what question your report is trying to answer.
 - Your research question is almost certainly about what factors influence your dependent variable (e.g., "What personal characteristics affect whether people think torture of suspected terrorists is at least sometimes justified?").
 - Briefly explain why your question is worth answering.
 - Hypotheses:
 - State and justify three (or more) hypotheses about factors that influence your dependent variable (1-3 sentences each).
 - Justifications can be very brief – they just need to show why you expect your hypothesis to be true (e.g., "Men will be more likely than women to believe that torture of suspected terrorists is at least sometimes justified. This is because men are more violent than women.").
 - Data:
 - Very briefly describe the dataset: who gathered the data, when, how, and what the sample size is. If you are using the birth control dataset you can use the codebook but make sure that you are not plagiarizing. If you use five words in a row from the codebook, then you need to put them in quotation marks and provide a citation. It's much better if you rewrite to paraphrase the information (and you'll still cite it).
 - State how you measure your variables.
 - Typically, start with the dependent variable, **which will almost certainly require giving the exact question wording**. Tell us what the values are.
 - Most independent variables will be clear enough that you do not need to provide the question wording but add it if necessary.
 - Be clear about any recoding you've done (e.g., combining values, reordering values, setting some values as missing). I do not need to know what the original coding was—leave that out of the text—but as an instructor, I want to be sure you did all your recodes correctly. In the appendix, show me a crosstab of the original and the recoded variable.

- Remember that you can only use nominal-, dummy-, and ordinal-level variables. If you want to use an interval-level variable (e.g., age or salary), then you need to recode it into an ordinal (or maybe a dummy) variable.
- Findings:
 - Provide pretty tables instead of just copying and pasting messy SPSS output. Do not include too many extraneous numbers (e.g., don't include all types of percentages in your tables if you only use one type of percentage). Add variable labels that make sense and look good. Try copying the SPSS table into Excel, throwing out all the columns/numbers you don't need, and getting rid of excessive digits before you copy into Word. See the SFdata table guidelines on D2L if you want great looking tables.
 - Start by describing the frequency distribution of your dependent variable (e.g., what percentage of respondents believe that torture of suspected terrorists is at least sometimes justified?). Do you need the frequency distribution table, or are the numbers simple enough that it's unnecessary? Would a graph be better than a table?
 - Briefly interpret your findings. That is, explain the relationships you find (or don't find) in clear language. You should use the appropriate percentages and percentage-point differences. Organize your discussion of findings around your hypotheses, probably discussing the relationships in the same order as the hypotheses.
 - In general, you'll want to include your tables just before/after the text where you discuss them.
 - Think of this as a short paper rather than a homework assignment or exam. That is, explain the findings in simple, straightforward language that someone not paid to do so would be willing to read.
 - Do not repeat the same paragraph over and over, just changing the variables and numbers. Try to make it interesting to read.
 - Throughout your paper, strive to use active voice verbs. (Avoid the passive voice. Try not to start sentences with "There is." If you are editing your text and notice that you have started sentences with phrases like "It could be argued that" or "It is believed that" or "Extant research indicates that," delete everything up through the "that.")
- Conclusion:
 - Very briefly review whether the data supported your hypotheses or not. Have you made progress in answering your research question?
 - If appropriate, does this suggest additional steps you should take? Do you see weaknesses/limitations in your answers? You might discuss additional variables you wish you could have included. You might discuss problems you see with the variables: Did they bias responses by asking a question in a particular way? Will everyone who ever smoked pot admit that they did so? Will other characteristics affect their willingness to do so? If so, is that likely to bias your findings?
- Appendix:

- Add this section of additional tables, etc. at the very end, starting on a new page.
- Add frequency distributions for each of the variables you used.
- If you recode any variables, add crosstabs between the old and new versions.
- Present the relevant portions of the crosstabs. (Don't include pieces that are meaningless to you.) I want to see the original SPSS tables in the appendix, as well as your cleaned-up version in the paper.

If you have trouble with formatting in Word, you may submit your appendix as a separate file. The text should probably come out to about 3 double-spaced pages (1 in. margins, 12 pt. font) without the tables.

Other tips for writing with this:

- (a) use the word 'percent' rather than the symbol '%'
- (b) don't begin a sentence with a numeral--spell out the number or use an introductory word
- (c) don't spell out numbers larger than ten
- (d) don't spell out any number with a decimal point or with the word 'percent'
- (e) don't use more than one digit past the decimal point in any percentage
- (f) remember that, for almost all uses in papers for me, 'affect' is a verb and 'effect' is a noun
- (g) consider using Grammarly